Concentration for Infinitely Divisible Vectors with Independent Components

C. Houdré* P. Reynaud-Bouret†

February 2, 2008

Abstract

For various classes of Lipschitz functions we provide dimension free concentration inequalities for infinitely divisible random vectors with independent components and finite exponential moments.

The purpose of this note is to further visit the concentration phenomenon for infinitely divisible vectors with independent components in an attempt to obtain dimension free concentration.

Let $X \sim ID(\gamma, 0, \nu)$ be an infinitely divisible (i.d.) vector (without Gaussian component) in $\mathbb{R}^d$, and with characteristic function $\varphi(t) = \mathbb{E}e^{i\langle t,X\rangle}$, $t \in \mathbb{R}^d$ (throughout, $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product in $\mathbb{R}^d$, while $\|\cdot\|$ is the corresponding Euclidean norm). As well known,

$$\varphi(t) = \exp\left\{ i\langle t,\gamma\rangle + \int_{\mathbb{R}^d} (e^{i\langle t,u\rangle} - 1 - i\langle t,u\rangle \mathbf{1}_{\|u\|\le 1})\nu(du) \right\}, \tag{1}$$

where $\gamma \in \mathbb{R}^d$ and where $\nu \not\equiv 0$ (the Lévy measure) is a positive Borel measure on $\mathbb{R}^d$, without atom at the origin and such that $\int_{\mathbb{R}^d}(1 \wedge \|u\|^2)\nu(du) < +\infty$. As also well known, X has independent components if and only if ν is supported on the axes of $\mathbb{R}^d$, i.e.,

$$\nu(dx_1,\dots,dx_d) = \sum_{k=1}^{d} \delta_0(dx_1)\cdots\delta_0(dx_{k-1})\tilde{\nu}_k(dx_k)\delta_0(dx_{k+1})\cdots\delta_0(dx_d). \tag{2}$$

*Laboratoire d'Analyse et de Mathématiques Appliquées, CNRS UMR 8050, Université Paris XII, 94010 Créteil Cedex, France and School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA.

†DMA Ecole Normale Supérieure, 45 rue d'Ulm 75230 Paris Cedex 05, France and School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA.

Moreover, the independent components of X have same law if and only if, the one dimensional Lévy measures $\tilde{\nu}_k$ are the same measure denoted by $\tilde{\nu}$.

Below, and throughout, by f Lipschitz with constant a we mean that $|f(x) - f(y)| \leq a\|x - y\|$, for all $x, y \in \mathbb{R}^d$ (the Lipschitz convention stated in [H] also applies). Let us start by recalling the following simple lemma which will be crucial to our approach [HPAS].

Lemma 1 *Let $X \sim ID(\gamma, 0, \nu)$ be such that $\mathbb{E}\|X\|^2 < +\infty$. Let $f, g : \mathbb{R}^d \to \mathbb{R}$ be Lipschitz functions. Then,*

$$\begin{aligned}\mathbb{E}f(X)g(X) &- \mathbb{E}f(X)\mathbb{E}g(X) \\ &= \int_0^1 \mathbb{E}_z \left[\int_{\mathbb{R}^d} (f(U+u) - f(U))(g(V+u) - g(V))\nu(du) \right] dz, \end{aligned} \tag{3}$$

where the expectation $\mathbb{E}_z$ is with respect to the i.d. vector, (U, V) in $\mathbb{R}^{2d}$ of parameter (γ, γ) and with Lévy measure $z\nu_1 + (1 - z)\nu_0$, $0 \leq z \leq 1$. The measure ν_0 is given by

$$\nu_0(du, dv) = \nu(du)\delta_0(dv) + \delta_0(du)\nu(dv), u, v \in \mathbb{R}^d,$$

while ν_1 is the measure ν supported on the main diagonal of $\mathbb{R}^d$.

An important feature of the representation (3) is the fact that the first marginal of (U, V) is X and so is its second marginal.

With the above framework and denoting by $e_1, e_2, \ldots, e_d$, the canonical basis of $\mathbb{R}^d$, we first prove:

Theorem 1 *Let $X = (X_1, \ldots, X_d) \sim ID(\gamma, 0, \nu)$ have independent components and be such that $\mathbb{E}e^{t\|X\|} < +\infty$, for some $t > 0$. Let $f : \mathbb{R}^d \to \mathbb{R}$, and let there exist $b_k \in \mathbb{R}$, $k = 1, \ldots, d$, such that $|f(x + ue_k) - f(x)| \leq b_k|u|$, for all $u \in \mathbb{R}$, $x \in \mathbb{R}^d$. Let*

$$h_f(t) = \sup_{x \in \mathbb{R}^d} \sum_{k=1}^{d} \int_{\mathbb{R}} |f(x + ue_k) - f(x)|^2 \frac{e^{tb_k|u|} - 1}{b_k|u|} \tilde{\nu}_k(du),\ 0 \leq t < M,$$

where $M = \sup\left\{t > 0 : \forall\ k = 1, \ldots, d, \mathbb{E}e^{tb_k|X_k|} < +\infty\right\}$. Then

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \geq x) \leq e^{-\int_0^x h_f^{-1}(s)ds}, \tag{4}$$

for all $0 < x < h_f^{-1}(M^-)$.

Proof. The proof is akin to proofs given in [H], and the above result complements the results there. First, by independence,

$$\begin{aligned}
C &= \{t > 0 : \forall\, k = 1, \ldots, d, \mathbb{E}e^{tb_k|X_k|} < +\infty\} \\
&= \left\{t > 0 : \forall\, k = 1, \ldots, d, \int_{|u|>1} e^{tb_k|u|}\tilde{\nu}_k(du) < +\infty\right\}.
\end{aligned}$$

Next, we apply the covariance representation (3) to f satisfying the above hypotheses and moreover assumed to be bounded and such that $\mathbb{E}f = 0$. Thus,

$$\begin{aligned}
\mathbb{E}fe^{tf} &= \int_0^1 \mathbb{E}_z\left[e^{tf(V)}\sum_{k=1}^d \int_{\mathbb{R}} (f(U+ue_k) - f(U))(e^{t(f(V+ue_k)-f(V))} - 1)\tilde{\nu}_k(du)\right] dz \\
&\le \int_0^1 \mathbb{E}_z\left[e^{tf(V)}\sum_{k=1}^d \int_{\mathbb{R}} |f(U+ue_k) - f(U)||f(V+ue_k) - f(V)|\,\frac{e^{tb_k|u|}-1}{b_k|u|}\,\tilde{\nu}_k(du)\right] dz \\
&\le \int_0^1 \mathbb{E}_z\left[e^{tf(V)}\sum_{k=1}^d \int_{\mathbb{R}} \frac{|f(U+ue_k) - f(U)|^2 + |f(V+ue_k) - f(V)|^2}{2}\left(\frac{e^{tb_k|u|}-1}{b_k|u|}\right)\tilde{\nu}_k(du)\right] dz \\
&\le h_f(t)\mathbb{E}\left[e^{tf}\right],
\end{aligned}$$

where we have used the "marginal property" mentioned above and since $h_f(t)$ is well defined for $0 \le t < M$. Integrating this last inequality, applied to $f - \mathbb{E}f$, leads to

$$\mathbb{E}e^{t(f-\mathbb{E}f)} \le e^{\int_0^t h_f(s)ds}, \qquad 0 \le t < M, \tag{5}$$

for all f bounded satisfying the hypotheses of the theorem. Fatou's lemma allows to remove the boundedness assumption in (5).

To obtain the tail inequality (4), the Bienaymé-Chebyshev inequality gives

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \ge x) \le \exp\left(-\sup_{0<t<M}\left(tx - \int_0^t h_f(s)ds\right)\right) = e^{-\int_0^x h_f^{-1}(s)ds},$$

by standard arguments, e.g., see [H]. ■

Theorem 1 is a bit formal, and we are now going to provide various cases where more concrete estimates are possible. Our first corollary, of Bennett–Prokhorov type, improves the constants in a result of [H]. If the components of X are iid Poisson random variables, then

(6) recovers also a result obtained by Bobkov and Ledoux [BL2] via modified log-Sobolev inequalities. This corollary is optimal in the one dimensional case, but suboptimal in the multidimensional one (see Corollary 5, for a more dimension free result)

Corollary 1 *Assume the hypotheses of the previous theorem. Moreover, let ν have bounded support with*

$$R_k = \inf\{\rho > 0 : \tilde{\nu}_k(|x| > \rho) = 0\}.$$

Set $bR = \max_{1\le k\le d} b_k R_k$ and set $\bar{a}^2 = \sup_{x\in\mathbb{R}^d} \sum_{k=1}^d \int_{|u|\le R_k} |f(x+ue_k)-f(x)|^2 \tilde{\nu}_k(du)$. Then, for all $x > 0$,

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \ge x) \le e^{-\frac{\bar{a}^2}{b^2R^2}\ell\left(\frac{bRx}{\bar{a}^2}\right)}, \tag{6}$$

where $\ell(u) = (1+u)\log(1+u) - u$, $u > 0$.

Proof. It is enough to note that $M = +\infty$ and that

$$h_f(s) \le \bar{a}^2 \left(\frac{e^{sbR} - 1}{bR}\right).$$

Integrating the reciprocal, gives $\frac{-x}{bR} + \left(\frac{x}{bR} + \frac{\bar{a}^2}{b^2R^2}\right) \log\left(1 + \frac{bRx}{\bar{a}^2}\right)$. ■

Let us now give a result which holds for Lévy measures with unbounded support, giving a Bernstein type inequality.

Corollary 2 *Assume the hypotheses of Theorem 1. Let $X \sim ID(\gamma, 0, \nu)$ have iid components and let f be such that $\tilde{a}^2 = \sup_{\substack{x\in\mathbb{R}^d \\ u\in\mathbb{R}, u\neq 0}} \sum_{k=1}^d \frac{|f(x+ue_k)-f(x)|^2}{|u|^2}$, and $b = \max_{1\le k\le d} b_k$ are finite. Then for all $0 < xb/\tilde{a}^2 < h^{-1}(M^-)$,*

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \ge x) \le \exp^{\left(-\frac{\tilde{a}^2}{b^2}\int_0^{xb/\tilde{a}^2} h^{-1}(s)ds\right)}, \tag{7}$$

where $h(s) = \int_{\mathbb{R}} |u|(e^{s|u|} - 1)\tilde{\nu}(du)$.

Moreover, if there exist $C > 0$ and $V^2 > 0$ such that

$$\int_{\mathbb{R}} |u|^n \tilde{\nu}(du) \le \frac{C^{n-2}n!}{2} V^2, \qquad \forall n \ge 2, \tag{8}$$

then, for all $x > 0$,

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \geq x) \leq e^{-\frac{\tilde{a}^2 V^2}{b^2 C^2}\ell\left(\frac{bCx}{\tilde{a}^2 V^2}\right)}, \tag{9}$$

where now $\ell(u) = (1+u) - \sqrt{1+2u}, u > 0$.

Proof. Again, we just need to bound h_f of Theorem 1. For (7), we bound h_f by

$$h_f(s) \leq \frac{\tilde{a}^2}{b}\int_{\mathbb{R}} u^2(e^{sbu} - 1)\tilde{\nu}(du),$$

and the result follows.

The condition (8) implies exponential moments, for $0 < t < \frac{1}{bC}$. Moreover,

$$\begin{aligned}
h_f(t) &\leq \sup_x \int_{\mathbb{R}} \sum_{k=1}^{d} \frac{|f(x+ue_k) - f(x)|^2}{|u|^2}\frac{|u|(e^{tb|u|} - 1)}{b}\,\tilde{\nu}(du) \\
&\leq \frac{\tilde{a}^2}{b}\frac{V^2}{2C}\sum_{k=2}^{\infty} k(tbC)^{k-1} \\
&= \frac{\tilde{a}^2}{b}\frac{V^2}{2C}\left(\frac{1}{(1-tbC)^2} - 1\right),
\end{aligned}$$

using (8). Integrating its reciprocal, we get

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \geq x) \leq e^{-\frac{\tilde{a}^2 V^2}{b^2 C^2}\left(1+\frac{bCu}{\tilde{a}^2 V^2} - \sqrt{1+\frac{2bCu}{\tilde{a}^2 V^2}}\right)}.$$

■

Remark 1 (i) An instance of the potential suboptimality of the previous results is the case of the (symmetric) exponential measure. Indeed, if $X_1, \ldots, X_d$ are iid with density $2^{-1}e^{-|x|}$, then the exponent in (9) or in (7) is of order $\min\left(\frac{x}{b}, \frac{x^2}{\tilde{a}^2}\right)$, while an inequality of Talagrand [T] asserts that the order $\min\left(\frac{x}{b}, \frac{x^2}{a^2}\right)$, where

$$a^2 = \sup_{x,u\in\mathbb{R}^d} \frac{|f(x+u) - f(x)|^2}{\|u\|^2}, \tag{10}$$

holds true. Clearly, $a^2 \leq \tilde{a}^2 \leq db^2$. It is then clear that (9) or (7) are optimal for linear functions, or infimum like above but not for the Euclidean norms. Actually, the example of

the Euclidean norm, i.e., $f(x) = \|x\|$, for which $a^2 = 1$, $b^2 = 1$, while $\tilde{a}^2 = db^2 = d$ shows that the concentration inequalities obtained that way are not dimension free. This is after all quite natural since Theorem 1 is really a result about "ℓ_1-Lipschitz" functions. We could also (to mimic the Lévy measure of the exponential law) replace in (8) $n!$ by $(n-1)!$, but the corresponding estimate will not be dimension free either. Our next result will show that for the Euclidean norm, a better estimation of (4) leads to dimension free concentration (see also Theorem 2 and Theorem 3).

(ii) Here is, however, an example of a class of function for which we can exactly get a dimension-free exponential inequality. Let X be an iid vector as in Corollary 2, for which $\tilde{\nu}$ has a support included in $\mathbb{R}^+$. Let f be defined by

$$f(x) = \inf_{\alpha \in A} f_\alpha(x),$$

where the $\{f_\alpha, \alpha \in A\}$ are non decreasing coordinates by coordinates and such that for all $1 \le k \le d$ and $\alpha \in A$, there exists a constant $b_{\alpha,k}$ such that

$$\forall u \in \mathbb{R}^+, \forall x \in \mathbb{R}^d, 0 \le f_\alpha(x + ue_k) - f_\alpha(x) \le b_{\alpha,k} u.$$

Then we can apply Corollary 2 to f. First let us remark that the supremum in $\tilde{a}$ can be taken for $u > 0$, and that

$$\tilde{a}^2 \le \sup_{x \in \mathbb{R}^d, u > 0} \sum_{k=1}^{d} \frac{|f_{\hat{\alpha}}(x + ue_k) - f_{\hat{\alpha}}(x)|^2}{u^2},$$

where $\hat{\alpha}$ is the index where $\inf_{\alpha \in A} f_\alpha(x)$ is achieved. Therefore, $\tilde{a}^2 \le \sup_{\alpha \in A} \sum_{k=1}^{d} b_{\alpha,k}^2$. On the other hand, one also has $b = \sup_{1 \le k \le d, \alpha \in A} b_{\alpha,k}$. This gives f, Lipschitz, nonlinear, with $\tilde{a} < \sqrt{d} b$. In particular, $f(X) = \inf_{1 \le k \le d} X_k$ verifies these conditions with $\tilde{a} = a = b = 1$, and Corollary 2 implies a complete dimension-free exponential inequality for the deviations of f above its mean.

Corollary 3 *Let $X \sim ID(\gamma, 0, \nu)$ have independent components and be such that $\mathbb{E}e^{t\|X\|} < +\infty$, for some $t > 0$. Let $M = \sup\{t > 0 : \forall\ k = 1, \dots, k,\ \mathbb{E}e^{t|X_k|} < +\infty\}$. Let $\varepsilon > 0$. Then, for all $0 < x < h(M^-)$*

$$\mathbb{P}(\|X\| \ge (1+\varepsilon)\mathbb{E}\|X\| + x) \le e^{-\int_0^x h^{-1}(s) ds}, \tag{11}$$

and

$$\mathbb{P}(\|X\| \leq (1-\varepsilon)\mathbb{E}\|X\| - x) \leq e^{-\int_0^x h^{-1}(s)ds}, \tag{12}$$

where the (dimension free) function h is given by $h(t) = 8\max_{1\leq k\leq d}\int_{\mathbb{R}} |u|(e^{t|u|}-1)\tilde{\nu}_k(du) + \frac{2}{(\varepsilon\mathbb{E}\|X\|)^2}\sum_{k=1}^d \int_{\mathbb{R}} |u|^3(e^{t|u|}-1)\tilde{\nu}_k(du)$.

Proof. We apply Theorem 1 to $f(x) = (\|x\| - \varepsilon\mathbb{E}\|X\|)^+$. Let us compute the various parameters and integrals for this f. First, it is easily verified that for each k, $|f(x+ue_k) - f(x)| \leq |\|x+ue_k\| - \|x\||\,\mathbf{1}_{A_k}$, where $A_k = \{\|x+ue_k\| \geq \varepsilon\mathbb{E}\|X\| \text{ or } \|x\| \geq \varepsilon\mathbb{E}\|X\|\}$, and where $\varepsilon > 0$. We then have

$$|f(x+ue_k) - f(x)| \leq \frac{|u(2x_k+u)|\mathbf{1}_{A_k}}{\|x+ue_k\| + \|x\|} \leq \frac{2|ux_k|}{\|x\|} + \frac{u^2}{\varepsilon\mathbb{E}\|X\|}, \tag{13}$$

where x_k is the kth coordinate of x. Moreover, since $|f(x+ue_k) - f(x)| \leq |u|$, we have

$$\begin{aligned}
&\sum_{k=1}^d \int_{\mathbb{R}} |f(x+ue_k) - f(x)|^2 \frac{e^{tb_k|u|}-1}{b_k|u|}\,\tilde{\nu}_k(du) \\
&\qquad \leq \sum_{k=1}^d \int_{\mathbb{R}} \left(8u^2\frac{|x_k|^2}{\|x\|^2} + \frac{2u^4}{(\varepsilon\mathbb{E}\|X\|)^2}\right)\left(\frac{e^{t|u|}-1}{|u|}\right)\tilde{\nu}_k(du).
\end{aligned}$$

Hence h_f in Theorem 1 is such that

$$h_f(t) \leq 8\max_{1\leq k\leq d}\int_{\mathbb{R}} |u|(e^{t|u|}-1)\tilde{\nu}_k(du) + \frac{2}{(\varepsilon\mathbb{E}\|X\|)^2}\sum_{k=1}^d \int_{\mathbb{R}} |u|^3(e^{t|u|}-1)\tilde{\nu}_k(du).$$

To finish the proof of (11) note that $\|X\| - \varepsilon\mathbb{E}\|X\| \leq (\|X\| - \varepsilon\mathbb{E}\|X\|)^+$ and that $\mathbb{E}(\|X\| - \varepsilon\mathbb{E}\|X\|)^+ \leq \mathbb{E}\|X\|$. To get the lower bound (12), just proceed as above but with the function $f(x) = -(\|x\| - \varepsilon\mathbb{E}\|X\|)^+$ and note that $(\|X\| - \varepsilon\mathbb{E}\|X\|)^+ \leq \|X\|$ and that $(1-\varepsilon)\mathbb{E}\|X\| \leq \mathbb{E}(\|X\| - \varepsilon\mathbb{E}\|X\|)^+$.

■

Remark 2 (i) The function h in the previous result is dimension free. Indeed,

$$h(t) \leq 8\max_{1\leq k\leq d}\int_{\mathbb{R}} |u|(e^{t|u|}-1)\tilde{\nu}_k(du) + \frac{2d}{(\varepsilon\mathbb{E}\|X\|)^2}\max_{1\leq k\leq d}\int_{\mathbb{R}} |u|^3(e^{t|u|}-1)\tilde{\nu}_k(du),$$

but,

$$d \min_{1\le k\le d} (\mathbb{E}|X_k|)^2 \le (\mathbb{E}\|X\|)^2 \le d \max_{1\le k\le d} \mathbb{E}(X_k^2).$$

(ii) When the Lévy measure has bounded support, the previous result leads under the assumptions of Corollary 1 to

$$\mathbb{P}(\|X\| \ge (1+\varepsilon)\mathbb{E}\|X\| + x) \le e^{-\left(\frac{x}{R}+\frac{V^2}{R^2}\right)\log\left(1+\frac{xR}{V^2}\right)+\frac{x}{R}}, \tag{14}$$

where

$$V^2 = 8 \max_{1\le k\le d} \left(\int_{|u|\le R} u^2 \tilde{\nu}_k(du)\right) + \frac{2}{(\varepsilon\mathbb{E}\|X\|)^2} \sum_{k=1}^{d} \int_{|u|\le R} u^4 \tilde{\nu}_k(du), \tag{15}$$

does not depend on the dimension d. This implies that there exists a constant $C > 0$ independent of d such that

$$\mathbb{P}(\|X\| \ge (1+\varepsilon)\mathbb{E}\|X\| + x) \le e^{-C\min\left(\frac{x}{R}\log\left(\frac{xR}{V^2}\right),\frac{x^2}{V^2}\right)}, \tag{16}$$

for all $x > 0$. ((14) and (15) allow to improve, for the Euclidean norm, the range and the constants in the last theorem of [HM].)

A direct consequence of Corollary 1 in [H] is the fact that for X infinitely divisible in $\mathbb{R}^d$ with boundedly supported Lévy measure,

$$\mathbb{E}e^{\frac{\|X\|}{R}\log^+\left(\frac{\lambda\|X\|}{R}\right)} < +\infty, \tag{17}$$

for all $\lambda > 0$ such that $\lambda V^2/R^2 < 1/e$, where $V^2 = \int_{\|u\|\le R} \|u\|^2 \nu(du)$. Although tight (take a one dimensional Poisson random variable with mean one), (17) is not optimal. Indeed a result of Rosiński [R] asserts that (for i.d. vectors in Banach spaces) (17) holds under the tighter condition $\lambda p_0 < 1/e$, where $p_0 = \nu(\|u\| = R)$. Similarly, another direct consequence of (14) is the following fact.

Corollary 4 *Let X be as in Corollary 3 above,*

$$\mathbb{E}e^{\frac{\|X\|}{R}\log^+\left(\frac{\lambda\|X\|}{R}\right)} < +\infty, \tag{18}$$

for all $\lambda > 0$ such that $\lambda V^2/R^2 < 1/e$, where now V^2 is given by (15).

Hence, for vectors with independent components, this last condition on λ is dimension free and in this sense, it improves on the general result obtained in [R]. Although dimension free, the condition $\lambda V^2/R^2 < 1/e$ (with V^2 as in (15)) is not optimal (again, take a one dimensional Poisson random variable with mean one). In view of [R], and say for X with iid components, one might wonder if $\lambda p_0 < 1/e$, where $p_0 = \tilde{\nu}(|u| = R)$ might be optimal.

The estimate (16) also improves a case of the exponential inequality derived for suprema of integrals with respect to a centered inhomogeneous Poisson process in [RB]. Let N be a Poisson process on $\mathbb{X}$ with intensity s with respect to μ. Let $\mathcal{P}$ be a partition of $\mathbb{X}$ and S the space of piecewise constant functions on $\mathcal{P}$. Let

$$\chi = \sup_{f \in S} \frac{\int_{\mathbb{X}} f \frac{dN - sd\mu}{\mu(\mathbb{X})}}{\sqrt{\int_{\mathbb{X}} f^2 \frac{d\mu}{\mu(\mathbb{X})}}}.$$

For this special choice of S, Proposition 9 of [RB] implies that, for all positive ε, there exists $C > 0$, such that

$$\forall x > 0, \mathbb{P}(\chi \geq (1+\varepsilon)\sqrt{\mathbb{E}\chi^2} + x) \leq \exp\left[-C \min\left((\sqrt{\eta\mu(\mathbb{X})})x, \frac{\mu(\mathbb{X})}{K} x^2\right)\right], \tag{19}$$

where $\eta = \inf_{I \in \mathcal{P}} \mu(I)$ and $K = \sup_{I \in \mathcal{P}} \frac{\int_I sd\mu}{\mu(I)}$.

But χ can be viewed as the Euclidean norm of the infinitely divisible vector

$$X = \left(\int_I \frac{dN - sd\mu}{\sqrt{\mu(I)\mu(\mathbb{X})}}\right)_{I \in \mathcal{P}},$$

with independent components. The Lévy measures of the components (see (2)) are given by

$$\tilde{\nu}_I = \left(\int_I sd\mu\right) \delta_{\frac{1}{\mu(I)\mu(\mathbb{X})}}.$$

Thus, we can apply (14) or (16) with $R = 1/\sqrt{\eta\mu(\mathbb{X})}$ and

$$V^2 = c(\varepsilon)\frac{K}{\mu(\mathbb{X})}.$$

Above, the constant $c(\varepsilon)$ does not depend on $\mathcal{P}$ or $\mu(\mathbb{X})$ as soon as $\eta > 1$ which is the interesting case where this type of inequality leads to adaptive estimator of the intensity s. We refer to [RB] for a complete description of this procedure. Therefore, (16) gives an extra

logarithmic factor with respect to (19) when S is a space of piecewise constant functions on a given partition. More precisely, for all positive ε, there exists $C > 0$, such that

$$\forall x > 0, \mathbb{P}(\chi \geq (1+\varepsilon)\mathbb{E}\chi + x) \leq \exp\left[-C \min\left(\sqrt{\eta\mu(\mathbb{X})}x \log\left(\frac{x\sqrt{\eta}\mu(\mathbb{X})^{3/2}}{K}\right), \frac{\mu(\mathbb{X})}{K}x^2\right)\right]. \tag{20}$$

(iii) If X has iid components $X_1, \ldots, X_d$ and if X_1 has an exponential distribution with density $2^{-1}e^{-|x|}$, $x \in \mathbb{R}$, and Lévy measure $|u|^{-1}e^{-|u|}$, $u \in \mathbb{R}, u \neq 0$, the previous result is a version of Talagrand's inequality (for norms). Indeed, in this case $M = 1$ and since $\mathbb{E}(|X_k|) = 1$, we obtain for all $0 \leq t < 1$,

$$\begin{aligned}
h(t) &= 8\int_{\mathbb{R}} |u|(e^{t|u|}-1)\frac{e^{-|u|}}{|u|}du + \frac{2}{\varepsilon^2(\mathbb{E}\|X\|)^2}d\int_{\mathbb{R}} |u|^3(e^{t|u|}-1)\frac{e^{-|u|}}{|u|}du, \\
&\leq 16\int_0^\infty (e^{tu}-1)e^{-u}du + \frac{4}{\varepsilon^2}\int_0^\infty u^2(e^{tu}-1)e^{-u}du, \\
&\leq 16\left(\frac{1}{1-t}-1\right) + \frac{8}{\varepsilon^2}\left(\frac{1}{(1-t)^3}-1\right), \\
&\leq \left(16+\frac{8}{\varepsilon^2}\right)\left(\frac{1}{(1-t)^3}-1\right).
\end{aligned}$$

This leads to

$$\forall x > 0, \quad \mathbb{P}(\|X\| \geq (1+\varepsilon)\mathbb{E}\|X\| + x) \leq e^{-x-\frac{3}{2}\left(16+\frac{8}{\varepsilon^2}\right)\left(1-\left(1+\frac{x}{\left(16+\frac{8}{\varepsilon^2}\right)}\right)^{2/3}\right)},$$

which implies that

$$\forall x > 0, \quad \mathbb{P}(\|X\| \geq (1+\varepsilon)\mathbb{E}\|X\| + x) \leq \exp\left(-\frac{x^2}{6(16+\frac{8}{\varepsilon^2})+4x}\right).$$

If one is only interested in the order of magnitude of the deviation of $\|X\|$, this is completely equivalent to Talagrand's inequality applied to the Euclidean norm, since (forgetting the constants and the dependency in ε) the exponent above is of order $-\min(x, x^2)$. However, one may want to get the exact upper deviation of $\|X\|$ from its mean (and not a constant times its mean). To see the difference, let us look at the reverse form :

$$\forall u > 0, \quad \mathbb{P}(\|X\| \geq (1+\varepsilon)\mathbb{E}\|X\| + \frac{4\sqrt{3u}}{\varepsilon} + 4\sqrt{6u} + 4u) \leq e^{-u}.$$

We can then minimize in ε and get:

$$\forall u > 0, \quad \mathbb{P}(\|X\| \geq \mathbb{E}\|X\| + 4(3u)^{1/4}(\mathbb{E}\|X\|)^{1/2} + 4\sqrt{6u} + 4u) \leq e^{-u}.$$

But $\mathbb{E}\|X\|$ grows like $\sqrt{d}$. So for d large, the quadratic term disappears and this is equivalent to

$$\forall x > 0, \quad \mathbb{P}(\|X\| \geq \mathbb{E}\|X\| + x) \leq e^{-C\min(x, \frac{x^4}{d})},$$

for some constant C. Hence, for d large, our method loses the quadratic behavior with respect to [T].

Our result is more restrictive than Talagrand's since it is only proved for norms rather than for arbitrary Lipschitz functions, and cannot give the exact order for the upper deviations from the mean but it is also more general since valid for any i.d. law with finite exponential moments (note too that the Lipschitz image of the exponential can be a bounded random variable and thus not i.d. and that not any i.d. variable with exponential moment is a Lipschitz image of the exponential variable.

(iv) A generalization of Corollary 3 to $\|X\|_A = \sqrt{X^*AX}$, where $A = (a_{j,k})$ is a symmetric positive definite matrix, is also possible. It is sufficient to remark that for all x in $\mathbb{R}^d$, $\|Ax\|^2 \leq \lambda_{\max}\|x\|_A^2$ where $\lambda_{\max}$ is the largest eigenvalue of A. Then we can apply Theorem 1 to $f(x) = (\|x\|_A - \varepsilon\mathbb{E}\|X\|_A)^+$, noticing that $b_k = \lambda_{\max}$ works and that

$$h_f(t) \leq 8\lambda_{\max} \max_{1\leq k\leq d} \int_{\mathbb{R}} |u|(e^{t|u|} - 1)\tilde{\nu}_k(du) + \frac{2}{(\varepsilon\mathbb{E}\|X\|_A)^2} \sum_{k=1}^{d} a_{k,k}^2 \int_{\mathbb{R}} |u|^3(e^{t|u|} - 1)\tilde{\nu}_k(du).$$

This upper bound is dimension free since $\|x\|_A^2 \geq \lambda_{\min}\|x\|^2$ where $\lambda_{\min}$ is the smallest eigenvalue of A.

We can in fact prove a result true for every Lipschitz function, by using the same type of method.

Theorem 2 *Let X be as in* Theorem 1. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be Lipschitz, with constant a. Then,*

$$\mathbb{P}\left(f(X) \geq \mathbb{E}f(X) + a\sqrt{2\sum_{k=1}^{d} \operatorname{Var} X_k} + ax \right) \leq e^{-\int_0^x h^{-1}(s)ds},$$

for all $0 < x < h(M^-)$, *where now*

$$h(t) = 8 \max_{1 \le k \le d} \int_{\mathbb{R}} |u|(e^{t|u|} - 1)\tilde{\nu}_k(du) + \frac{2}{\sum_{k=1}^d \operatorname{Var} X_k} \sum_{k=1}^d \int_{\mathbb{R}} |u|^3 (e^{t|u|} - 1)\tilde{\nu}_k(du).$$

Proof. We apply Theorem 1 to $\phi(X) = \sqrt{\mathbb{E}_Y \|X - Y\|^2}$, where Y is a vector such that $\mathbb{E}_Y \|Y\|^2 < +\infty$ and independent of X. As, $\sqrt{\mathbb{E}_Y \| \cdot \|^2}$ is a norm (for vectors depending on Y), we have

$$|\phi(X + ue_k) - \phi(X)| \le \sqrt{\mathbb{E}_Y \|X + ue_k - Y - (X - Y)\|^2} \le |u|,$$

Thus $b_k = 1$, for all $k = 1, \ldots, d$. Also,

$$|\phi(X + ue_k) - \phi(X)|^2 = \left(\frac{\mathbb{E}_Y (2u(X_k - Y_k) + u^2)}{\sqrt{\mathbb{E}_Y \|X - Y\|^2} + \sqrt{\mathbb{E}_Y \|X + ue_k - Y\|^2}} \right)^2 . \tag{21}$$

Note that $\phi(X) \ge \sqrt{\sum_{k=1}^d \operatorname{Var} Y_k}$. Hence, the right hand side of (21) is dominated by

$$\frac{8u^2 \mathbb{E}_Y (X_k - Y_k)^2}{\mathbb{E}_Y \|X - Y\|^2} + \frac{2u^4}{\sum_{k=1}^d \operatorname{Var} Y_k} . \tag{22}$$

We then see (using (22)) that the function h_ϕ in Theorem 1 is such that

$$h_\phi(t) \le 8 \max_{1 \le k \le n} \int_{\mathbb{R}} |u|(e^{t|u|} - 1)\tilde{\nu}_k(du) + \frac{2 \sum_{k=1}^d \int_{\mathbb{R}} |u|^3 (e^{t|u|} - 1)\tilde{\nu}_k(du)}{\sum_{k=1}^d \operatorname{Var} Y_k}.$$

Returning to f, and taking for Y an independent copy of X, we get

$$\mathbb{E}\phi(X) \le \sqrt{\mathbb{E}_X \mathbb{E}_Y \|X - Y\|^2} = \sqrt{2 \sum_{k=1}^d \operatorname{Var} X_k}.$$

Moreover, $|f(X) - \mathbb{E}f(X)| \le a\phi(X)$. These last two estimates finally give

$$\mathbb{P}\left(f(X) \ge \mathbb{E}f(X) + a\sqrt{2 \sum_{k=1}^d \operatorname{Var} X_k} + ax \right) \le \exp\left(- \int_0^x h^{-1}(t)dt \right).$$

■

Remark 3 The above result gives a dimension-free exponential rate of decay for the deviations of $f(X)$ above its mean plus $a\sqrt{d}$, up to some constants. For the exponential distribution, Theorem 2 does not give an exponential rate with two speeds, one using b (defined in Corollary 2) and the other using a (defined by (10)). This cannot be seen either in Corollary 3, since for the Euclidean norm $a = b$. But one can combine Theorem 1 and Theorem 2 together. For iid variables with exponential symmetric distributions, this gives a partial version of Talagrand's result [T]. First, we look at the deviation of f above $m = \mathbb{E}f(X) + 2a\sqrt{d}$. As $\mathbb{E}f(X) \leq m$, from Corollary 2 there exists some absolute constant c_1 such that

$$\mathbb{P}\left(f(X) \geq m + x\right) \leq \exp\left(-c_1 \min\left(\frac{x}{b}, \frac{x^2}{\tilde{a}^2}\right)\right),$$

and from Theorem 2 there exists some absolute constant c_2 such that

$$\mathbb{P}\left(f(X) \geq m + x\right) \leq \exp\left(-c_2 \min\left(\frac{x}{a}, \frac{x^2}{a^2}\right)\right).$$

This implies that there exists some absolute constant c_3 such that

$$\mathbb{P}\left(f(X) \geq m + x\right) \leq \exp(-c_3 g(x))$$

where

$$\begin{aligned} g(x) &= \frac{x^2}{a^2}, \text{ for } 0 \leq x \leq a, \\ &= \frac{x}{a}, \text{ for } a \leq x \leq \frac{\tilde{a}^2}{a}, \\ &= \frac{x^2}{\tilde{a}^2}, \text{ for } \frac{\tilde{a}^2}{a} \leq x \leq \frac{\tilde{a}^2}{b}, \\ &= \frac{x}{b}, \text{ for } \frac{\tilde{a}^2}{b} \leq x. \end{aligned}$$

Thus we recover Talagrand's result for small and large x. In the middle, we have intermediate rate. If $a = b = 1$ (as for the Euclidean norm) or if $a = \tilde{a}$ (as for linear functionals), we recover exactly Talagrand's rate on the whole real line. For the deviation with respect to $\mathbb{E}f(X)$ and not $\mathbb{E}f(X) + 2a\sqrt{d}$, the previous rates become worse, but sometimes improve the rate given by Corollary 2 for some special parts of the real line.

The next result is an easy consequence of Theorem 2 by applying the same methods as in the proof Corollary 1. Combined with Corollary 1, it will give dimension free rates in e^{-x^2/a^2}, for x small above $a\sqrt{d}$, and of order $e^{-\frac{x}{bR}\log x}$, for x large.

Corollary 5 *Let X be as in Theorem 1. Moreover, let ν have bounded support with*

$$R = \max_{1\leq k\leq d} \inf\{\rho > 0 : \tilde{\nu}_k(|x| > \rho) = 0\}.$$

Let f be a Lipschitz function with constant a.

Then, for all $x > 0$,

$$\mathbb{P}\left(f(X) \geq \mathbb{E}f(X) + a\sqrt{2\sum_{k=1}^{d} \operatorname{Var} X_k} + ax \right) \leq e^{-\frac{v^2}{R^2}\ell\left(\frac{Rx}{v^2}\right)}, \tag{23}$$

where $\ell(u) = (1+u)\log(1+u) - u$, $u > 0$ and

$$v^2 = 8 \max_{1\leq k\leq d} \int_{\mathbb{R}} |u|^2 \tilde{\nu}_k(du) + \frac{2}{\sum_{k=1}^{d} \operatorname{Var} X_k} \sum_{k=1}^{d} \int_{\mathbb{R}} |u|^4 \tilde{\nu}_k(du).$$

Remark 4 (i) The above improves Corollary 1 as one can see on a vector of iid Poisson variables with parameter 1. The quantity $\bar{a}$ appearing in Corollary 1 is then equal to $\tilde{a}$ appearing in Corollary 2 and is of order $\sqrt{d}$, while Corollary 5 gives a dimension-free exponential rate of decay for the deviations of f above $\mathbb{E}f(X) + a\sqrt{2d}$.

(ii) A natural question is then to know whether or not the above result is a consequence of, or implies, a result of Bobkov and Ledoux [BL1] which asserts that a Poincaré inequality does imply Talagrand's. This is not the case. First, a uniform random variable on $[0,1]$ satisfies a Poincaré inequality but is not infinitely divisible. Second, a Poisson random variable has finite exponential moments, is infinitely divisible but does not satisfy a Poincaré inequality. However, Corollary 5 combined with Corollary 1 gives dimension free rates in e^{-x^2/a^2}, for x small above $a\sqrt{d} + \mathbb{E}f(X)$ and of order $e^{-\frac{x}{bR}\log x}$, for x large. This is almost a dimension free inequality with two rates except that f has to exceed $\mathbb{E}f(X) + a\sqrt{d}$ and not just $\mathbb{E}f(X)$ and that there are smaller rates for intermediate x.

Of course, we would like a result using only a and b for every Lipschitz functions to exactly recover the exponential case. In particular, even if f has to exceed a multiple of

$a\sqrt{d}$, we would like to improve the rates obtained in Remark 3 when $a \le x \le \tilde{a}^2/b$ and $a > b$. The next two results give some further knowledge in this direction. The first one deals with concave functions and so also leads to a left tail inequality for the Euclidean norm.

Corollary 6 *Let X be as in* Theorem 1, *let $f : \mathbb{R}^d \to \mathbb{R}$ be concave and let* $\tilde{b}_k = \left|\mathbb{E}\,\frac{\partial f(X)}{\partial x_k}\right|$, $k = 1, \dots, d$. *Let* $M = \sup\left\{t > 0 : \forall\ k = 1, \dots, d, \mathbb{E}e^{t\tilde{b}_k|X_k|} < +\infty\right\}$. *Let* $\mathrm{Cov}(X, \nabla f(X)) = \mathbb{E}\langle X, \nabla f(X)\rangle - \langle \mathbb{E}X, \mathbb{E}\nabla f(X)\rangle$. *Then,*

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \ge -\mathrm{Cov}(X, \nabla f(X)) + x) \le e^{-\int_0^x h^{-1}(s)ds}, \tag{24}$$

for all $0 < x < h(M^-)$, *where* h *is given by* $h(t) = \sum_{k=1}^d \int_{\mathbb{R}} \tilde{b}_k|u|(e^{t\tilde{b}_k|u|}-1)\tilde{\nu}_k(du)$, $0 < t < M$.

Proof. Since f is concave, and if Y is an independent copy of X,

$$f(X) - \mathbb{E}f(X) \le \mathbb{E}_Y(\langle X - Y, \nabla f(Y)\rangle) := \phi(X).$$

We then apply Theorem 1 to ϕ. Indeed,

$$\phi(X + ue_k) - \phi(X) = u\mathbb{E}\frac{\partial f(Y)}{\partial x_k}.$$

Hence, $h_\phi(t) = \sum_{k=1}^d \int_{\mathbb{R}} \tilde{b}_k|u|(e^{t\tilde{b}_k|u|} - 1)\tilde{\nu}_k(du)$, and the result follows. ■

Remark 5 Above, if $|f(x) - f(y)|^2 \le a^2\|x - y\|^2$ we get:

$$\begin{aligned} h_\phi(t) &\le \sum_{k=1}^d \tilde{b}_k^2 \int_{\mathbb{R}} |u|^2 \frac{(e^{t\tilde{b}_k|u|} - 1)}{\tilde{b}_k|u|}\,\tilde{\nu}_k(du) \\ &\le \max_{1\le k\le n} \int_{\mathbb{R}} u^2 \frac{(e^{t\tilde{b}_k|u|} - 1)}{\tilde{b}_k|u|}\,\tilde{\nu}_k(du) \sum_{k=1}^d \left(\mathbb{E}\frac{\partial f(X)}{\partial x_k}\right)^2 \\ &\le a^2 \max_{1\le k\le d} \int_{\mathbb{R}} |u| \left(\frac{e^{t\tilde{b}_k|u|} - 1}{\tilde{b}_k}\right) \tilde{\nu}_k(du). \end{aligned}$$

Moreover,

$$0 \le \mathbb{E}\phi(X) = -\mathrm{Cov}(X, \nabla f(X)) \le \mathbb{E}\left(\|X - \mathbb{E}X\|\|\nabla f(X)\|\right) \le a\mathbb{E}\|X - \mathbb{E}X\|.$$

Combining these two facts we see that (24) becomes

$$\mathbb{P}\left(f(X) - \mathbb{E}f(X) \geq a\mathbb{E}\|X - \mathbb{E}X\| + x\right) \leq e^{-\int_0^x h^{-1}(t)dt}, \tag{25}$$

where now $h(t) = a^2 \max_{1\leq k\leq 1} \int_{\mathbb{R}} |u| \frac{(e^{\tilde{b}_k|u|}-1)}{\tilde{b}_k} \tilde{\nu}_k(du)$. This last inequality is once again dimension free. In the particular case, $f(X) = -\|X\|$ $(a = b_k = 1)$, we get $-\mathrm{Cov}(X, \nabla f(X)) = \mathbb{E}\|X\| - \langle \mathbb{E}X, \mathbb{E}\left(X/\|X\|\right)\rangle$ and (24) becomes

$$\mathbb{P}\left(-\|X\| \geq -\langle \mathbb{E}X, \mathbb{E}\left(X/\|X\|\right)\rangle + x\right) \leq e^{-\int_0^x h^{-1}(t)dt}, \tag{26}$$

with $h(t) = \max_{k=1,\ldots,d} \int_{\mathbb{R}} |u|(e^{t|u|} - 1)\tilde{\nu}_k(du)$. However, the inequality (26) does not present any interest when $\mathbb{E}X = 0$.

The second result deals with general Lipschitz functions, gives exponential inequalities using a and b and allows us to improve the rates, in the exponential case, when $a\sqrt{d} < \tilde{a}^2/b$ and $a > b$.

Theorem 3 *Let X be as in Theorem 1. Let f be a Lipschitz function with constant a, and let $b_k \in \mathbb{R}$, $k = 1, \ldots, d$, such that $|f(x + ue_k) - f(x)| \leq b_k|u|$, for all $u \in \mathbb{R}$, $x \in \mathbb{R}^d$. Let $\varepsilon > 0$. Then, for all $0 < x < h^{-1}(M)$*

$$\mathbb{P}\left(f(X) \geq f(0) + a\mathbb{E}\left([\|X\| - \varepsilon]_+\right) + a\varepsilon + x\right) \leq e^{-\int_0^x h^{-1}(s)ds}, \tag{27}$$

where

$$h(t) = 2a\sqrt{\sum_{k=1}^d \left(\int_{\mathbb{R}} |u|(e^{tb_k|u|} - 1)\tilde{\nu}_k(du)\right)^2} + \frac{a}{\varepsilon}\sum_{k=1}^d \int_{\mathbb{R}} |u|^2(e^{tb_k|u|} - 1)\tilde{\nu}_k(du),$$

for all $0 \leq t < M$.

Proof. Let $M > t \geq 0$. First, we have

$$\mathbb{E}\left(\left[f(X) - f(0) - a\varepsilon - a\mathbb{E}\left([\|X\| - \varepsilon]_+\right)\right] e^{tf(X)}\right) \leq \mathrm{Cov}(g(X), e^{tf(X)}),$$

where $g(X) = a[\|X\| - \varepsilon]_+$. By using equation (3), we get that

$$\begin{aligned}\operatorname{Cov}(g(X), e^{tf(X)}) &= \int_0^1 \mathbb{E}_z\left[e^{tf(V)} \sum_{k=1}^d \int_{\mathbb{R}} (g(U+ue_k) - g(U))(e^{t(f(V+ue_k)-f(V))} - 1)\tilde{\nu}_k(du)\right] dz \\ &\leq \int_0^1 \mathbb{E}_z \left[e^{tf(V)} \sum_{k=1}^d \int_{\mathbb{R}} |g(U+ue_k) - g(U)||f(V+ue_k) - f(V)| \frac{e^{tb_k|u|} - 1}{b_k|u|} \tilde{\nu}_k(du)\right] dz.\end{aligned}$$

By using the computations done in the proof of Corollary 3, we know that

$$|g(U+ue_k) - g(U)| \leq 2a|u|\frac{|U_k|}{\|U\|} + \frac{au^2}{\varepsilon}.$$

Let us define A and B by

$$A = \int_0^1 \mathbb{E}_z \left[e^{tf(V)} \sum_{k=1}^d \int_{\mathbb{R}} 2a\frac{|U_k|}{\|U\|}|f(V+ue_k) - f(V)| \frac{e^{tb_k|u|} - 1}{b_k} \tilde{\nu}_k(du)\right] dz,$$

$$B = \mathbb{E} \left[e^{tf(V)} \sum_{k=1}^d \int_{\mathbb{R}} \frac{a|u|}{\varepsilon}|f(V+ue_k) - f(V)| \frac{e^{tb_k|u|} - 1}{b_k} \tilde{\nu}_k(du)\right].$$

Then we obtain that

$$\mathbb{E}\left([f(X) - f(0) - a\varepsilon - a\mathbb{E}\left([\|X\| - \varepsilon]_+\right)] e^{tf(X)}\right) \leq A + B.$$

We can bound A by

$$\begin{aligned}A &\leq \int_0^1 \mathbb{E}_z \left[e^{tf(V)} \sum_{k=1}^d \int_{\mathbb{R}} 2a\frac{|U_k|}{\|U\|}|u|\left(e^{tb_k|u|} - 1\right) \tilde{\nu}_k(du)\right] dz \\ &\leq 2a\sqrt{\sum_{k=1}^d \left(\int_{\mathbb{R}} |u|(e^{tb_k|u|} - 1)\tilde{\nu}_k(du)\right)^2}\, \mathbb{E}(e^{tf(X)}).\end{aligned}$$

Similarly, we get the following upper bound for B:

$$B \leq \frac{a}{\varepsilon}\left(\sum_{k=1}^d \int_{\mathbb{R}} |u|^2(e^{tb_k|u|} - 1)\tilde{\nu}_k(du)\right) \mathbb{E}(e^{tf(X)}).$$

It remains to use the classic integration/maximisation method to conclude the proof. ■

Remark 6 (i) Comparing Theorem 2 and Theorem 3, we see that $f(0)$ is replacing $\mathbb{E}f(X)$. This is not a problem since (see [MS, Appendix V]) one can pass from the former to the later up to some multiplicative constant. In fact, it was already possible to derive directly Theorem 2 with $f(0)$ instead of $\mathbb{E}f(X)$.

(ii) If X is a vector of iid variables with density $2^{-1}e^{-|x|}$, and if $b = \max_k b_k$, after computations (similar to the ones given in Remark 2 (iii)), we obtain that for every Lipschitz function f, for all $x, \varepsilon > 0$,

$$\mathbb{P}(f(X) \geq f(0) + a\sqrt{d} + a\varepsilon + x) \leq e^{-c\min(\frac{x}{b}, \frac{x^2}{ab\sqrt{d}+2ab\frac{d}{\varepsilon}})}, \tag{28}$$

for some absolute constant c. The reverse form of this last inequality is more practical in order to better understand the various orders of magnitude: for all $x, \varepsilon > 0$,

$$\mathbb{P}(f(X) \geq f(0) + a\sqrt{d} + a\varepsilon + \square\sqrt{\frac{abdx}{\varepsilon}} + \square\sqrt{ab\sqrt{d}x} + \square bx) \leq e^{-x}, \tag{29}$$

where the $\square$ are known absolute constants. So first if one take $\varepsilon = \delta\sqrt{d}$ then this implies

$$\mathbb{P}\left(f(X) \geq f(0) + (1+2\delta)a\sqrt{d} + \left(\square + \frac{\square}{\delta^2}\right) bx\right) \leq e^{-x}.$$

Thus, once f has exceeded $f(0)$ plus a multiple, as close to 1 as we want, of $a\sqrt{d}$ the behavior is linear, and the slope is b up to some multiplicative constant, increasing as δ tends to 0. This improves the results of Remark 3 for the exponential case when $\tilde{a}^2/b >> a\sqrt{d}$ and $\tilde{a} > a > b$ since now the linear rate x/b is true on a larger interval. This partially recovers Corollary 3 since $\sqrt{x} \leq \sqrt{d} + x$, for all $d \geq 1$ and $x > 0$.

One can also optimize in ε, getting that for all positive x

$$\mathbb{P}(f(X) \geq f(0) + a\sqrt{d} + \square a^{2/3}b^{1/3}d^{1/3}x^{1/3} + \square a^{1/2}b^{1/2}d^{1/4}x^{1/2} + \square bx) \leq e^{-x}.$$

(iii) This result also improves the rates for iid Poisson variables with parameter 1. When Corollary 5 gives the rate $\exp(-C_1\frac{x}{a}\log(\frac{x}{a}))$ for the deviations above $Ef + a\sqrt{d} + x$ for sufficiently large x, Theorem 3 gives $\exp(-C_1\frac{x}{b}\log(\frac{x}{a\sqrt{(d)}}))$ for $x > C_3 a\sqrt{d}$ which is better than Corollary 1 and Corollary 5 as soon as $b << a$ and $a\sqrt{d} << \tilde{a}^2/b$.

(iv) More generally, if one is interested in Lipschitz function of i.d. vectors with independent components and Lévy measure with bounded support, the equivalent of Corollary

5 can be obtained by applying Theorem 3 to Lévy measures with bounded support. Similarly, the equivalent of Corollary 4 for f, Lipschitz function with constant a (b being defined as usual) can also be obtained. One straightforward application is then to say that $|f(X) - f(0)| \leq a\|X\|$, giving :

$$\mathbb{E}e^{\frac{f(X)}{aR}\log^{+}(\frac{\lambda f(X)}{aR})} < \infty,$$

for all $\lambda > 0$ such that $\lambda V^2/R^2 < 1/e$, where V^2 is given by (15). But one may wonder if the above remains true with b instead of a, i.e., the ℓ^1-Lipschitz constant. By applying Theorem 3 with $\varepsilon = \sqrt{d}$, it follows that

$$\mathbb{E}e^{\frac{f(X)}{bR}\log^{+}\left(\frac{\lambda f(X)}{bR}\right)} < \infty, \tag{30}$$

for all $\lambda > 0$ such that $\lambda \frac{aV^2}{bR} < 1/e$, where this time

$$V^2 = 3\sqrt{d}\left(\max_{1\leq k\leq d}\left(\int_{\mathbb{R}}|u|\tilde{\nu}_k(du)\right) \vee \max_{1\leq k\leq d}\left(\int_{\mathbb{R}}|u|^2\tilde{\nu}_k(du)\right)\right).$$

As V^2 is not dimension free, this is not as sharp as Corollary 4 for the Euclidean norm, but it is sharper than the results of [R] since, in that case, V^2 would be of order d. It also implies with b instead of a the following result

$$\mathbb{E}e^{\frac{f(X)}{A}\log^{+}(\frac{f(X)}{bR})} < \infty,$$

for all $A > bR$, which is a complete dimension free result and which can be of interest if $a >> b$.

The various results presented here for vectors with finite exponential moment as well as the general methodology presented in [HM] delineate quite well the concentration phenomenon for infinitely divisible vectors. Nevertheless, and say, for iid components, it will be interesting to prove versions of Theorem 2 or of Theorem 3 for the deviations of an arbitrary Lipschitz function above its mean and not just above its mean plus $a\sqrt{d}$, up to a constant. Such a possible extension would then give, when combined with Theorem 1 a dimension-free exponential inequality with two rates rather than one, and as such would then give us a pretty complete understanding of this topic.

References

[BL1] Bobkov, S. G. and Ledoux, M. (1997). Poincaré inequalities and Talagrand's concentration phenomenon for the exponential measure. *Probab. Theory Related Fields* **107** 383–400.

[BL2] Bobkov, S. G. and Ledoux, M. (1998). On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.* **156** 347–365.

[HPAS] Houdré, C., Pérez-Abreu, V. and Surgailis, D. (1998). Interpolation, correlation identities and inequalities for infinitely divisible variables. *J. Fourier Anal. Appl.* **4** 651–668.

[H] Houdré, C. (2002). Remarks on deviation inequalities for functions of infinitely divisible random vectors. *Ann. Probab.* **30** 1223–1237.

[HM] Houdré, C. and Marchal P. (2002). On the concentration of measure phenomenon for stable and related random vectors. To appear: *Ann. Probab.*

[MS] Milman V.D. and Schechtman G. (1986) Asymptotic theory of finite-dimensional normed spaces. *Lecture Notes in Math.* **1200** Springer, Berlin.

[RB] Reynaud-Bouret, P. (2003). Adaptive estimation of the intensity of inhomogeneous Poisson processes via concentration inequalities. *Probab. Theory Related Fields* **126** 103–153.

[R] Rosiński, J. Remarks on strong exponential integrability of vector–valued random series and triangular arrays. *Ann. Prob.* **23** (1996), 464–473.

[T] Talagrand, M. (1991). A new isoperimetric inequality for product measure, and the concentration of measure phenomenon. Israel Seminar (GAFA). *Lecture Notes in Math.* **1469** 91–124. Springer, Berlin.